



**You have downloaded a document from
RE-BUS
repository of the University of Silesia in Katowice**

Title: Traduction automatique des langues - esperances et enjeux des outlis appliques

Author: Aleksandra Żłobińska-Nowak

Citation style: Żłobińska-Nowak Aleksandra. (2013). Traduction automatique des langues - esperances et enjeux des outlis appliques. "Neophilologica" (T. 25 (2013), s. 56-66).



Uznanie autorstwa - Użycie niekomercyjne - Bez utworów zależnych Polska - Licencja ta zezwala na rozpowszechnianie, przedstawianie i wykonywanie utworu jedynie w celach niekomercyjnych oraz pod warunkiem zachowania go w oryginalnej postaci (nie tworzenia utworów zależnych).



UNIwersYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego

Aleksandra Żłobińska-Nowak

*Université de Silésie
Katowice*

Traduction automatique des langues — espérances et enjeux des outils appliqués

Abstract

This paper deals with a short study of automatic translation and natural language processing in contemporary linguistics. The author presents the main reasons of their conception as well as their beginnings and expectations in the historical perspective. Also the difficulties in the domain of automatic translation like polysemy, ambiguity or choice of the formalisms appropriate to the description and the proposals of their resolution are focused on in the paper. In this context, the article examines the problem of word sense disambiguation according to the object oriented approach proposed by W. Banyś. The author seeks to demonstrate the differences in the choice of the linguistic means in French and Polish on the analyzed levels and the influence of occurrence in this description. The aim of this paper is also to indicate some possibilities to communicate our human linguistic abilities to the computer to ensure the correct machine translation.

Keywords

Automatic translation, natural language processing, polysemy, ambiguity, word sense disambiguation, object oriented approach, occurrence.

Cette communication se veut une brève étude de la situation de la traduction automatique et du traitement des langues dans la linguistique contemporaine. L'objectif est de répondre aux questions suivantes : quels sont les problèmes auxquels se heurte la discipline de la traduction automatique, comment communiquer différentes connaissances à la machine que nous, en tant qu'humains, possédons, développer et mener à bien ses capacités et comment assurer l'opérationnalité de l'approche à objets pour apprendre à l'ordinateur à maîtriser les langues soumises à la traduction et à s'orienter dans un espace interlinguistique où la langue n'est plus une barrière.

Nous allons commencer cette présentation par une courte esquisse concernant le domaine de la traduction automatique ainsi que le traitement des langues dans

le cadre de la linguistique contemporaine pour pouvoir en sortir les principes de la conception de ce type de traduction qui suscite aujourd'hui tant d'intérêt dans les analyses de la langue.

Définissons brièvement tout d'abord les termes de « traitement automatique des langues » et de « traduction automatique ». Le premier désigne un domaine qui vise à concevoir les logiciels qui pourraient traiter automatiquement, sans intervention du facteur humain ou avec son intervention partielle, des données linguistiques, exprimées dans une langue naturelle. Les données prises en compte peuvent être de différentes sortes (C. Fuchs, 1993 : 7—8) : textes écrits, textes oraux, phrases, énoncés, groupes de mots ou mots isolés. Le traitement équivaut à une modification d'un objet d'entrée en un objet de sortie. Les techniques et les outils appliqués dans le traitement automatique de langues peuvent être linguistiques, formels et informatiques. Le caractère linguistique de ces outils-là est lié à la description détaillée de l'objet traitée — une langue, suivant des procédures formelles précises employées en vue de leur expression dans des formalismes implémentables en machine (ordinateur) qui emploie, lui, certaines techniques d'ordre informatique. Pour appliquer donc un traitement à une langue naturelle il est indispensable de découvrir les principes régulateurs de constitution de chaque donnée provenant de la langue et de les décrire de manière opératoire qui inclurait les formes correspondant à des sens et serait soumise à un ensemble de règles explicitables.

Pour ce qui est de la traduction automatique, elle consiste à l'examen détaillé d'un texte, sa préparation à un traitement automatique et, en résultat, la récupération de sa traduction en une deuxième langue.

Bien entendu, les systèmes purement automatiques ne permettent pas de choix interactif et n'admettent aucune intervention humaine. Même s'ils deviennent de plus en plus fréquents dans le cadre de l'informatique actuelle, ils laissent toujours à désirer.

Il existe quelques types de systèmes d'une popularité importante dans la traduction, à savoir :

- MAHT (*Machine Aided Human Translation*) où la traduction est assurée par l'homme qui se sert de l'ordinateur en vue de la facilitation et l'accélération de ce processus. Ce type de système fournit au traducteur les outils tels que les dictionnaires de langue, terminologiques, d'expressions et de locutions, de synonymes, la fonction de vérification orthographique et grammaticale dans une autre langue qui rendent possible le développement de l'aspect plus créatif dépendant du savoir propre du traducteur. Dans la traduction du type MAHT le traducteur assume le rôle le plus important alors que l'ordinateur n'est que son outil.
- HAMT (*Human Aided Machine Translation*) qui exécute une traduction assistée par l'homme. Le texte source peut subir des modifications avant, en train ou après la traduction faite par l'ordinateur, par exemple dans le cas où il est nécessaire de préparer une forme spéciale du document, facile à être traitée

par une machine, d'introduire dans le texte source des informations qui rendraient possible l'analyse. Ainsi, peuvent être traduites, souvent avec succès, les notices des médicaments, les manuels et la documentation technique en plusieurs langues. Nous pouvons citer ici le fameux système TAUM-METEO créé à l'Université de Montréal et destiné à traduire les informations météorologiques de l'anglais vers le français et inversement pour les services météorologiques du gouvernement canadien. Les données introduites avaient le caractère spécialisé sans être un texte ordinaire en langue naturelle.

- FAMT (*Fully Automated Machine Translation*) est une traduction effectuée sans aucune aide de l'homme. Le logiciel prépare, à la base d'un texte source, un texte dans la langue cible. Dans cette opération, le résultat n'est pas toujours d'une qualité satisfaisante, néanmoins peut être utile dans certains objectifs, tels que : parcourir les sites Internet ou lire les mails là où nous avons besoin urgent d'une information tandis qu'une traduction correcte et entièrement acceptable n'est pas indispensable. Il est à remarquer également que ce type de traduction peut s'avérer utile aux utilisateurs dont la connaissance de la langue cible permet la compréhension du texte généré par la machine et sa correction, si tel est le besoin.
- Certaines sources distinguent encore un groupe supplémentaire de programmes à l'acronyme FAHQMT (*Fully Automated High Quality Machine Translation*) qui sont censés fournir une traduction très exacte et exhaustive dans des domaines linguistiques restreints (S. Kozłowski, 2004 : 67—68).

Il faut souligner que cette distinction n'est pas toujours adoptée par tous les linguistes et informaticiens qui parfois la réduisent en deux branches TA (*traduction automatique* non assistée par aucun traducteur humain) et TAO (*traduction assistée par ordinateur* où le traducteur humain peut dialoguer avec la machine au sujet des propositions de traduction qu'elle lui offre) (A.-M. Löffler-Laurian, 1996 : 7).

Les outils fournis par la machine qui traduit un texte, et que le traducteur humain peut modifier, sont fréquemment intégrés dans un éditeur de texte et cette technologie porte le nom de CAT (*Computer Aided Translation*). Nous disposons actuellement de plusieurs logiciels qui procèdent de cette manière, p.ex. WordFast, SDLTrados, MemoQ, DéjàVu.

Dans le traitement automatique des langues, les domaines les plus anciens sont le traitement de la parole et la traduction. La traduction automatique était l'un des objectifs les plus recherchés dans l'ensemble des outils possibles qu'offraient les ordinateurs depuis la Seconde Guerre mondiale. Nous pouvons recenser aujourd'hui trois principales étapes dans l'histoire de la TA :

1. Fin des années 40 jusqu'au rapport ALPAC en 1965 où l'on observe un mélange d'émotions autour du sujet de la TA. Les grandes idées enthousiastes des années 40 sont liées avec les techniques du déchiffrement cryptographique pouvant être utilisées dans la traduction automatique des textes et introduites dans cet objectif par Warren Weaver. Le projet suscitait un grand intérêt surtout dans le

contexte économique-politique concernant les États-Unis et l'URSS et pour cela les langues traitées se limitaient à l'anglais et au russe. L'an 1952 c'est la date de la première conférence sur la traduction automatique au M.I.T. C'est à cette période-là que les chercheurs doivent faire face à une désillusion des projets envisagés due à un manque de réflexion sur les langues soumises à la traduction. Les systèmes élaborés étaient dépourvus de représentation syntaxique ne proposant que des traductions mot à mot ce qui engendraient des échecs. Le fameux rapport du comité ALPAC de 1965 (*Automatic Language Processing Advisory Committee*) en évaluant le progrès en domaine de la traduction automatique constate que les recherches menées ne sont pas rentables pour l'État Américain et réduit les moyens accordés aux chercheurs qui doivent se tourner vers la linguistique informatique.

2. Dix années suivantes (1965—1975) sont marquées par une stagnation dans le domaine cependant quelques travaux se poursuivent, à signaler, à titre d'exemple les systèmes METAL, SYSTRAN ou TAUM. Les recherches effectuées se concentrent sur quelques aspects négligés avant tels que : une approche indirecte de la traduction reposant sur les modules distincts d'analyse de la langue-source et de génération de la langue-cible, une séparation entre les connaissances linguistiques (grammaire, dictionnaires) et la partie logicielle (programmes), une traduction faite à un niveau plus profond passant par une phase syntaxique facile à s'enrichir d'éléments de nature sémantique, une prise en compte des objectifs plus restreints qui consistent à envisager un domaine délimité avec un lexique et une syntaxe répétitifs (p.ex. les bulletins météo dans le système TAUM-METEO), une diversification des langues traduites.
3. À partir de 1975 sont repris les travaux dans le domaine de la traduction automatique et une commercialisation de leurs produits liée avec une demande considérable du marché. C'est aussi une période du développement des outils informatiques. Les systèmes de traduction proposés deviennent plus sophistiqués et font preuve des capacités élargies et évoluées des langages informatiques de l'époque. On arrive à une pluralité de langues pouvant être traitées dans le processus de la traduction (cf. p.ex. P. Bouillon, 1998 ; C. Fuchs, 1993).

Quelles sont les données linguistiques qu'un programme idéal devrait inclure dans sa version informatique pour pouvoir assurer une traduction automatique de langue ?

Ces informations devraient concerner la nature des mots analysés, leur prononciation, leur signification et leurs capacités de se combiner pour constituer des structures plus complexes y compris l'influence de leurs combinaisons sur le sens des phrases ainsi formées. Un tel programme devrait également rendre compte des connaissances générales sur le monde et les contextes d'utilisation des textes (cf. p.ex. P. Bouillon, 1998 ; C. Fuchs, 1993).

Les informations d'ordre phonétique et phonologique comptent dans le cas de la reconnaissance de la parole ainsi que dans sa synthèse. Il est question d'appren-

dre à une machine à distinguer les sons et la manière dont les mots se réalisent en sons.

Premièrement, la machine est donc censée assurer la phonétisation d'un texte écrit pour arriver à une chaîne de phonèmes en fournissant des corrélats acoustiques à ces objets phoniques dans le processus de la synthèse qui consiste à créer de la parole artificielle à partir de n'importe quel texte, utile, entre autres, dans la vocalisation d'écrans informatiques pour les personnes aveugles (http://fr.wikipedia.org/wiki/Synth%C3%A8se_vocale).

Deuxièmement, dans le cas de la reconnaissance automatique de la parole la machine doit disposer des techniques qui permettent d'analyser la parole captée pour la transcrire ensuite sous la forme d'un texte exploitable (http://fr.wikipedia.org/wiki/Reconnaissance_vocale).

Les connaissances à caractère morphologique sont censées répondre à la question comment sont construits les mots à la base des morphèmes, leurs unités minimales de signification. Il est donc question des mots ou des affixes qui s'ajoutent à ces mots-là. Dans cette branche linguistique du traitement automatique nous pouvons énumérer trois catégories :

- la morphologie flexionnelle centrée sur la combinaison des mots et des affixes sans influence sur le changement de la catégorie syntaxique (p.ex. *maison* — *maisons*),
- la morphologie dérivationnelle centrée sur la formation des mots nouveaux où les affixes changent de catégorie syntaxique du mot auquel ils se joignent (p.ex. *réparer* — *réparable*),
- la morphologie compositionnelle centrée sur le processus de la formation des mots nouveaux à partir de mots isolés (p.ex. *porte-manteau*).

Dans l'analyse morphologique on établit des listes ordonnées d'unités élémentaires issues de la segmentation du texte, ainsi, par exemple, dans le cas de la phrase *Jean a construit une baraque* nous pouvons observer les segments suivants :

1. *Jean*
2. *a construit*
3. *une*
4. *baraque*

où chaque élément correspondra à une forme répertoriée dans le système. Dans le cas des formes variées nous retrouverons une forme lemmatisée appropriée p.ex. pour *a construit*, la forme lemmatisée sera le verbe à l'infinitif *construire* avec les informations morpho-syntaxiques du type : verbe, passé composé, indicatif, 3^{ème} personne, singulier.

L'analyse syntaxique est censée fournir une représentation structurelle appropriée à une combinaison de mots en phrases, doit décrire comment les mots

forment des phrases syntaxiquement correctes et quelles sont les erreurs inacceptables au niveau de l'ordre des unités composantes (p.ex. *Il aime son chat. *Chat son aime il.*). Dans la plupart des analyseurs syntaxiques la phrase jouit d'un rôle privilégié étant donné que les unités de taille supérieure ne se prêtent pas facilement à un traitement automatique. De plus, dans la tradition linguistique une phrase constitue un tout cohérent au sein duquel s'annoncent surtout les principaux phénomènes de prédication (l'interaction entre le sujet et le prédicat qui permet de faire ressortir leurs sens ce que nous allons aborder un peu plus loin dans cette présentation).

Il faudrait remarquer à cette occasion que c'étaient les ouvrages de Noam Chomsky de la fin des années 50 avec lesquels la pensée linguistique a été marquée par une série de notions nouvelles annonçant fortement les débuts de la linguistique informatique, Chomsky se vouant lui-même à une branche récemment développée à cette époque-là surtout dans le M.I.T., à savoir, l'étude des langages formels et la linguistique informatique, tout ceci dans le cadre de l'Intelligence Artificielle.

L'analyse sémantique adoptée par le système de traitement automatique repose fréquemment sur les représentations syntaxiques élaborées préalablement ou simultanément à cette analyse-là. Comme nous pouvons nous imaginer les difficultés que ce type de description soulève, les études se limitent dans la plupart des cas à des domaines restreints laissant de côté les essais de l'analyse complète et générale de la langue, indépendante d'une branche quelconque. Le sens d'une phrase analysée correspond au calcul des informations fournies par la description morphologique, syntaxique et lexicale. Le sens résulte donc à la fois du sens des mots (sémantique lexicale) et du sens des relations entre ses mots-là (sémantique grammaticale) (C. Fuchs, 1993 : 140).

Même si actuellement les capacités, la puissance et la mémoire des ordinateurs ne constitue plus un problème au développement de l'informatique, bien au contraire, dépasse même notre imagination, la langue naturelle soumise à l'analyse et au traitement effectués par une machine dévoile toujours des difficultés parfois impossibles à surmonter.

Tel est le cas du fameux phénomène de la polysémie caractérisant toute langue donnée. La polysémie est définie dans le dictionnaire *Petit Robert* comme « caractère d'un signe qui possède plusieurs contenus, plusieurs sens » tandis que la notion d'ambiguïté renvoie dans le même ouvrage au « caractère de ce qui est ambigu (qui présente deux ou plusieurs sens possibles ; dont l'interprétation est incertaine) dans le langage ». Dans le *Dictionnaire actuel de la langue française* la polysémie est définie en tant que « capacité qu'offre un mot ou un texte de présenter plusieurs significations différentes selon la perspective de lecture », l'ambiguïté étant un « caractère de ce qui est ambigu (qui peut avoir deux sens différents ; équivoque) ». Pour résumer, nous pouvons constater que la polysémie est une capacité des mots appartenant à un système alors que l'ambiguïté en est une conséquence activée dans l'emploi où un mot, un énoncé peuvent avoir différentes interprétations.

L'ambiguïté remet en cause alors le potentiel polysémique des unités lexicales. La polysémisation n'est pas un phénomène accidentel de la langue, elle est entièrement explicable au sein du système linguistique auquel les lexèmes appartiennent, elle devient même indispensable pour ne pas arriver à une langue monosémique où tout référent du monde réel aurait un autre signifiant. La polysémie dans la langue obéit à la loi de l'économie du langage étant donnée qu'elle réutilise plusieurs fois le même signe en faisant varier son signifié. Autrement encore c'est une mise en rapport d'un même signifiant avec plusieurs signifiés comme, par exemple, dans le cas du mot *rouge* : la couleur, le vin, la race, la colère, le communiste, ou bien les verbes à haute fréquence présentant les polysémies les plus riches, p.ex. 39 acceptions du verbe *aller*, 49 acceptions du verbe *mettre* indiquées par Littré (<http://littre.reverso.net/dictionnaire-francais/>).

Bien entendu, au cours de la communication les locuteurs ne doivent pas forcément se rendre compte de l'ambiguïté de la phrase, c'est le contexte qui indique le sens du message transmis. Le sens global de ce message-là permet donc d'éclaircir d'éventuelles ambiguïtés.

Quelles sont les ambiguïtés les plus fréquentes ?

Il existe des ambiguïtés au niveau du mot lui-même qualifiées de lexicales. Dans ce groupe nous pouvons énumérer :

- les ambiguïtés catégorielles liées aux plusieurs catégories syntaxiques concernant une seule unité lexicale p.ex. le mot *bouche* qui peut être un nom ou un verbe à la troisième personne du singulier du verbe *boucher* ;
- les ambiguïtés sémantiques qui résultent du sens polysémique du mot comme, p.ex., le mot *rouge* cité ci-dessus ou encore de l'homonymie, p.ex. *vol*.

Pour pouvoir désambiguïser les mots ambigus nous pouvons recourir à des indices syntaxiques, comme une préposition ou une construction particulière qui permettront de faire ressortir le sens du mot analysé en contexte, p.ex. *douter de qqch* vs *se douter de qqch*. Cependant les autres cas, comme le mot *vol* imposent une construction détaillée d'un champ sémantique créé à partir des mots appartenant à un domaine spécifique du vocabulaire. Une telle construction s'inscrit dans la lignée de recherches proposées dans les années 70 par M. Minsky, R.C. Schank et R.P. Abelson et portant sur les *frames* et *scripts*. En gros, les frames constituent des structures comportant toutes les informations sur les objets analysés tandis que les scripts (scénarios) sont censés présenter les séquences d'actions stéréotypées (pour les frames R.C. Schank, R.P. Abelson, 1977 et les scripts M. Minsky, 1975). Nous allons présenter encore dans ce qui suit une autre conception de désambiguïstation à partir des classes d'objets.

Il existe également des ambiguïtés syntaxiques et structurelles (P. Bouillon, 1998 : 21) liées aux constructions des mots. Pour qu'une phrase soit syntaxiquement ambiguë elle doit admettre plusieurs lectures dues à sa construction. Nous avons à faire à une ambiguïté catégorielle quand une phrase constitue un exemple de variation flexionnelle, p.ex. *Une plie la plie*.

Nous pouvons également avoir affaire aux ambiguïtés structurelles qui résultent du découpage de la phrase (p.ex., le rattachement d'un syntagme prépositionnel ou d'une proposition relative), comme dans la phrase : *Le jardin du domaine qui a été vendu lui plaisait beaucoup* (le jardin ou le domaine qui a été vendu lui plaisait ?).

Pour pouvoir résoudre le problème des ambiguïtés structurelles le système devrait inclure les procédures de reconnaissance syntaxique (comme le type de complément introduit par le verbe) et sémantique (comme le type de restriction de sélection).

La méthode suivante laquelle nous effectuons nos travaux propose de désambiguïser le sens des mots par l'intermédiaire des ensembles qui partagent les mêmes traits sémantiques et les mêmes propriétés syntaxiques (cf. p.ex. D. Le Pesant, M. Mathieu-Colas, 1998). L'approche orientée objets conçue par W. Banyś (2002, 2005) souligne, dans l'analyse du sens des mots, l'importance de ces ensembles portant le nom de classes d'objets.

Ce sont des groupes de mots apparentés sémantiquement dont la présence autour des unités analysées, comme des verbes polysémiques par exemple, permet de dévoiler leur sens concret provenant de l'influence mutuelle de ces éléments (cf. p.ex. les travaux de G. Gross, 1992, 1994, 1995, 1997, 1999). Les classes d'objets avec les constructions syntaxiques particulières des verbes fournissent ensemble des outils pour la désambiguïsation.

En analysant, par exemple, l'inventaire des équivalents recensés dans les dictionnaires français-polonais du verbe *sortir* nous y trouvons plusieurs possibilités de traduction de ce verbe-là, à savoir : *wyjsć* (*skądś, z czegoś*), *wyjechać*, *wydobyć* (*kogoś z czegoś, coś skądś*), *wystąpić* (*z czegoś*), *wykoleić się*, *wzejść*, *ukazać się*, *wydstać* (*się z czegoś, kogoś skądś, coś skądś*), *wybrnąć* (*z czegoś*), *oddalić się*, *odejść* (*od czegoś*), *odstąpić* (*od czegoś*), *zboczyć* (*z czegoś*), *zejść* (*z czegoś*), *wyniknąć*, *wywieść*, *pochodzić* (*skądś, od kogoś, czegoś*), *wyprowadzić*, *wyciągnąć* (*coś skądś, kogoś skądś*), *wyjąć*, *wyrzucić* (*kogoś skądś*), *wypuścić* (*na rynek*), et d'autres qui peuvent s'y ajouter comme résultat de l'opération de désambiguïsation détaillée à la base des classes d'objets : *wystawać* (*z czegoś*), *wypaść* (*z czegoś, coś*), *wydawać*, *być wydanym*, *powiedzieć*, *rzucić*, *wypłynąć* (*z czegoś*), *wylać*, *wyskoczyć* (*na czymś*), *wyrósć* (*na czymś*), *być widocznym*, *wyrzec się* (*czegoś, kogoś*), *porzucić*, *odbiec* (*od czegoś*), *zjechać*, *wynieść*, *wyrwać* (*z czegoś*) (A. Żłobińska-Nowak, 2008 : 117—118).

Nous pouvons remarquer que le choix d'équivalents est très riche, leur étendue d'emploi en polonais paraît parfois bien limitée et est loin d'avoir une portée si générale que le sens du verbe *sortir* français. Ceci dit, *sortir* est qualifié de verbe hautement polysémique néanmoins la présence des classes d'objets particulières dans les phrases construites à sa base permet de choisir le verbe approprié polonais souvent, beaucoup moins polysémique ou à un sens relativement pauvre (tel est le cas de *wykoleić się*).

À titre d'exemple prenons le sens polonais *wylać się* qui, selon *Słownik języka polskiego PWN* peut avoir trois acceptions : 1. «o płynach: wypłynąć z jakiegoś naczynia, zbiornika itp.» 2. «wydostać się skądś, zwykle w dużej ilości» 3. «o uczuciach: uzewnętrznić się». Le schéma syntaxico-sémantique français du verbe *sortir* qui est à l'origine de cette traduction polonaise montre deux classes d'objets spécifiques (A. Żłobińska-Nowak, 2008 : 123) :

X — [CONC <liquide / substance liquide> ; <boisson>] — **sortir** — *de* — Y — [CONC <récipient qui peut contenir un liquide ou une boisson>] — **wylać się z**
Une goutte de trop et l'eau sort du vase.

Ainsi, la construction des classes d'objets impose-t-elle la spécification des liquides et des récipients dans lesquels ils se trouvent avant d'effectuer le mouvement de l'intérieur vers l'extérieur. Ce passage est le seul lien, le sens dit prototypique, qui unit le verbe *sortir* avec l'équivalent polonais *wylać się*.

Un autre moyen de désambiguïsation qui va de pair avec des classes d'objets est la construction particulière de la phrase comportant le lexème désambiguïsé. Parfois sa spécificité est si importante que la recherche de l'équivalent dans la traduction ne laisse pas de doutes (A. Żłobińska-Nowak, 2008 : 124) :

X — [CONC <véhicule de traction>] — **sortir** — *de* — Y — [CONC <limite : rails>] — **wykoleić się**
La locomotive est sortie des rails.
Le tram de Nancy sort des rails dans un virage.

Dans le cas présenté ci-dessus la classe d'objets <véhicule de traction> est un ensemble relativement limité, qui, avec le complément d'objet indirect rempli par le seul lexème *rails*, permet au système de reconnaître facilement de quel équivalent il est question.

Comme le souligne C. Fuchs (1993 : 13—15) le traitement automatique des langues naturelles a des enjeux sociaux, économiques, politiques, idéologiques et culturels. Il est né comme résultat du développement progressif de l'informatique et du besoin de traiter une masse d'informations. Ces informations ne sont pas réservées uniquement à des domaines socio-économiques mais concernent largement la société actuelle qui se sert quotidiennement des outils informatiques dans le travail ainsi que dans la vie privée. Pour ce qui est de la traduction automatique, elle peut fournir des gains considérables dans le secteur économique. Il est à observer également que la surévaluation de l'anglais en tant que langue principale des logiciels avait pour effet le rejet d'autres langues nationales en imposant à l'utilisateur le système culturel américain.

L'industrialisation des langues naturelles minoritaires qui n'avaient pas pour autant de statut de langues dominantes peut donc s'avérer utile dans la commu-

nication, dans la transmission des aspects culturels ainsi que dans l'obtention des profits matériels dans de nombreux secteurs économiques.

Nous espérons que les approches à objets pourront jouer un rôle important dans ces recherches.

Références

- Banyś W., 2002a : « Bases de données lexicales électroniques — une approche orientée objets. Partie I : Questions de modularité ». *Neophilologica*, **15**, 7—29.
- Banyś W., 2002b : « Bases de données lexicales électroniques — une approche orientée objets. Partie II : Questions de description ». *Neophilologica*, **15**, 206—249.
- Banyś W., 2005 : « Désambiguïsation des sens des mots et représentation lexicale du monde ». *Neophilologica*, **17**, 57—76.
- Bouillon P., 1998 : *Traitement automatique des langues naturelles*. Paris — Bruxelles, Aupelf-Uref — Éditions Duculot.
- Chiss J.-L., Filliolet J., Maingueneau D., 2001 : *Introduction à la linguistique française*. T. 1 : *Notions fondamentales, phonétique, lexique*. Paris, Hachette Livre.
- Fuchs C., 1993 : *Linguistique et traitements automatiques des langues*. Paris, Hachette Supérieur.
- Fuchs C., 1996 : *Les ambiguïtés du français*. Paris, Ophrys.
- Gross G., 1992 : « Forme d'un dictionnaire électronique ». In : *La station de traduction de l'an 2000*. Presses de l'Université du Québec.
- Gross G., 1994 : « Classes d'objets et description des verbes ». *Langages*, **115**, 15—31.
- Gross G., 1995 : « Une sémantique nouvelle pour la traduction automatique. Les classes d'objets ». *La Tribune des industries de la langue et de l'information électronique. Perspectives*, numéro spécial (n^{os} 17—19) : *Traduction et traduction avec outils, le renouveau pour demain*, 16—19.
- Gross G., 1997 : « La grammaire, les dictionnaires et l'informatique ». In : *Les dictionnaires de langue française et l'informatique. Actes du Colloque « La Journée des dictionnaires »*. Université de Cergy-Pontoise, Centre de Recherche Texte / Histoire, 55—65.
- Gross G., 1999 : « Élaboration d'un dictionnaire électronique ». In : *Bulletin de la Société Linguistique de Paris*, T. XCIV, fasc. 1. Leuven : Peeters, 113—138.
- Kozłowski S., 2004 : *Co to jest tłumaczenie maszynowe?* Warszawa : Koło Filozoficzne Uniwersytetu Warszawskiego. Disponible sur : http://kf.mish.uw.edu.pl/kog/kog_seb.pdf (accessible : le 3 juillet 2013).
- Le Pesant D., Mathieu-Colas M., 1998 : « Introduction aux classes d'objets ». *Langages*, **131**, 6—33.
- Loffler-Laurian A.-M., 1996 : *La traduction automatique*. Villeneuve-d'Ascq, Presses Universitaires de Septentrion.
- Minsky M., 1975 : « A Framework for Representing Knowledge ». In : T.H. Winston, C. Brown, eds., 1979 : *Artificial Intelligence*. Massachussets, MIT Press.

- Minsky M., 1986 : *The Society of Mind*. New York, Simon & Schuster (trad. fr. : *La Société de l'esprit*. Paris, InterEditions, 1988).
- Schank R.C., Abelson R.P., 1977: *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.
- Żłobińska-Nowak A., 2008 : *Désambiguïsation des expressions lexicales des opérateurs de l'espace dans le cadre d'une approche orientée objets : les verbes de mouvements "monter" et "sortir" et leurs équivalents polonais*. Katowice, Wydawnictwo UŚ.

Dictionnaires de langue

- Dobrzyński J., Kaczuba I., Frosztęga B., 1991 : *Grand dictionnaire français-polonais*. T. 1—2. Warszawa, WP.
- Dictionnaire actuel de la langue française*, 1991. Paris, Librairies Flammarion et Quillet.
- Grand Larousse de la langue française en six volumes*, 1971. Direction L. Guilbert, R. Lagane, G. Niobey. Paris, Larousse.
- Larousse dictionnaire de français 35 000 mots*, 1986. Direction J. Dubois. Paris, Larousse.
- Le nouveau Petit Robert*, 1995. Paris, Dictionnaires Le Robert.
- Słownik języka polskiego PWN*. T. 1—3, wersja 1.0. Warszawa, Wydawnictwo Naukowe PWN.
- Dictionnaire de français „Littré”*. Disponible sur : <http://littrereverso.net/dictionnaire-francais/> (accessible : le 3 juillet 2013).